

LA-UR-21-24071

 $\label{lem:proved} \mbox{Approved for public release; distribution is unlimited.}$

Title: National Security Research Center

Author(s): Ali, Alee Rizwan

Intended for: presentations

Issued: 2021-04-27







NATIONAL SECURITY







MAbout Us



- Born from the need to make LANL
 Weapons Program archival material
 available to scientists and engineers
- Material is the end result of decades of consolidation of mini libraries and mini archives at LANL
- Latest consolidation brought together LANL's digital archives and physical archives
- Houses 75+ years of nuclear weapons research, designs, procedures, videos, photos, and other reports





The NSRC offers services similar to major university research libraries.





LIBRARY COMPARISON (BY HOLDINGS)

MAJOR NATIONAL **SECURITY LIBRARIES**

2.6M

1.5M

I.IM

NATIONAL **SECURITY** RESEARCH **CENTER**

AIR UNIVERSITY LIBRARY

OAK RIDGE **NAVAL POST** NATIONAL GRADUATE LABORATORY **SCHOOL** LIBRARY LIBRARY



MAJOR ARCHIVAL LIBRARIES



NATIONAL

SECURITY

RESEARCH

CENTER

9.6M

LBJ

PRESIDENTIAL

LIBRARY

5.3M

KENNEDY PRESIDENTIAL PRESIDENTIAL LIBRARY

HOOVER LIBRARY

4.8M





16.9M



13M



NATIONAL **SECURITY** RESEARCH CENTER

OF CHICAGO UNIVERSITY LIBRARY

COLUMBIA UNIVERSITY LIBRARY









NSRC's Collections



- Not a traditional archive.
- NSRC's holdings are active collections used daily by researchers at LANL and throughout the NSE.
- Contains *millions* of unique and historically significant items documenting the U.S. nuclear weapons design, development, testing, certification, and assessment activities.

SAMPLE OF NSRC HOLDINGS
Nuclear weapons designs
Assembly drawings
Nuclear test diagnostic data
Pre-shot reports
Post-shot reports and analyses
Supporting and confirmatory experiments
Subcritical experiments
Stockpile stewardship experiments
Milestone reports
Classified research journals
US/UK collaboration documents
Internal memoranda
Working group notes
Histories
Code manuals
Test problems

The vast majority of the material is not duplicated anywhere else in the NSE.



Sample of NSRC's Collections



-an	ш	
	ш	

COLLECTION	CONTENT
Rocky Flats Collection	 Detailed pit production, certification, and use information Technical reports describing efforts by researchers to develop new pit technology
Classified Reports Collection	 Studies of Soviet missile defense Effects of underwater blasts on submarine integrity Requirements for strategic stability Development of warheads for advanced delivery systems Special processes for pit welding
Top Secret Collection	From the intelligence community, Joint Chiefs of Staff, the Pentagon, and other government agencies
United Kingdom Collection	 7,000+ accountable docs Originally supported the 1958 Mutual Defense Agreement Relevant in weaponry design, naval nuclear propulsion, and nuclear threat reduction
Wellnitz Collection	 Legacy weapons engineering drawings, hydro test records, and material NSTEC (Nevada National Security Site) vault collection Old J-Division (field testing) collection New material from weapons physicists
Los Alamos Historical Collection	 Archival material related to the Manhattan Project Correspondences, films, reports, engineering drawings, lab notebooks









- Less than 10% of the physical collection is digitized.
- Decades-long project to digitize the following media types:
 - ---> Paper
 - ---> Engineering drawings
 - ---> Microfilm
 - ---> Microfiche
 - ---> Photos
 - --- Aperture cards

- ---> Video
- ---> Motion picture film
- --- Audio
- --- Notebooks
- --- Negatives
- --- Card catalog

The backlog in digitizing cannot be addressed with current processes.



Digitizing Efforts – Historical Rates







- ~1,300 videos per year
 - → Total collection: ~10,000 (6 videos per day)
 - → ~8 years to digitize
- ~15 microfilm rolls per year
 - → Total collection: ~30,000 (1,500 2,000 pages per film roll)
 - → ~2,000 years to digitize
- ~2,200 microfiche sheets per year
 - ---> Total collection: ~200,000 (10 20 pages per sheet)
 - ---> ~93 years to digitize
- ~52 "banker boxes" of paper digitized per year
 - ---> Total collection: ~40,000 (~2,000 pages per box)
 - → ~760 years to digitize

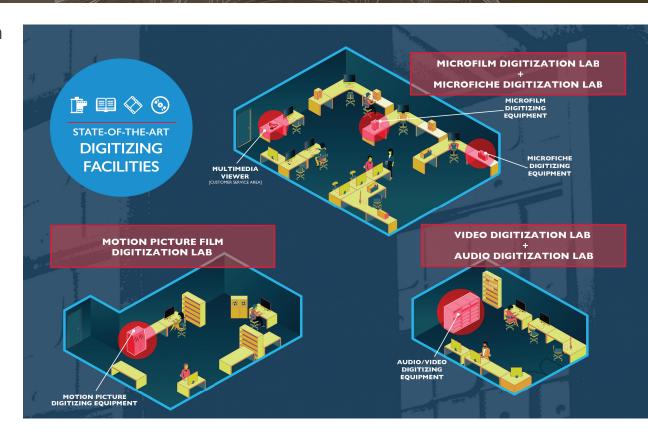
Digitization is limited by equipment, software, processes, training, and facilities.



Status of Digitization Services



- Seven new digitization labs in FY20 & FY21:
 - ---> Video Tape Digitization Lab
 - ---> Audio Tape Digitization Lab
 - ---> Microfilm Digitization Lab
 - ---> Microfiche Digitization Lab
 - ---> Motion Picture Film Digitization Lab
 - ---> Rocky Flats Digitization Lab
 - ---> Specialized Digitization Lab (paper, engineering drawings, photos, lab notebooks, aperture cards, etc.)
- Established digitization equipment certification program:
 - ---> All applicable staff to be certified on equipment and archival standards.



Refocus Attention on Highest-Priority Media



- Video (VHS, Betamax, etc.)
 - ---> Fastest deteriorating media in our collections
 - ---> Rate of digitization may not address deteriorating media
 - ---> Increased speed through number of digitization suites / FTEs
 - --> Estimated completion: Reduced from 8 years to 5 years

- Microfilm Of immediate relevance to Weapons Program
 - ---> New production-scale equipment and software
 - ---> Revamped processes
 - → Increase of personnel from 0.1 to 2
 - ---> Estimated completion: Reduced from 2,000 years to 16 years
 - Microfiche Of <u>immediate relevance</u> to Weapons Program
 - ---> New production-scale equipment and software
 - ---> Revamped processes
 - → Increase of personnel from 1 to 3
 - --> Estimated completion: Reduced from 93 years to 19 years

Video	Legacy	New
Quantity	10,000	
FTEs	1.0	1.5
Digitization suites	2	4
Rate per month	108	170
Years to complete	8	5

Microfilm	Legacy	New
Quantity	30,000	
FTEs	0.1	2
Digitization suites	1	3
Rate per month	1.3	160
Years to complete	2,000	16

Microfiche	Legacy	New
Quantity	200,000	
FTEs	1	3
Digitization suites	1	4
Rate per month	180	900
Years to complete	93	19



Cataloging / Indexing Efforts



4-1-

Digital Collections

- ---> Wide variety of metadata and indices ranging from nonexistent to rudimentary to robust
- ---> Search functions range from nonexistent to difficult-to-use

Physical Collections

- ---> Poor to nonexistent metadata
- ---> Limited indices
- ---> Classified Reports Collection has an old-fashioned card catalog

Cataloging rate:

- ---> Currently using *fully manual* process
- ---> Process takes 10-30 minutes per document
- ---> Total number of *digitized files growing* rapidly

Cataloging snapshot of one digitized collection		
Quantity	2.4 million	
FTEs	1.5	
Rate per month	486	
Years to complete 412		

The backlog in cataloging/indexing cannot be met with a manual approach.

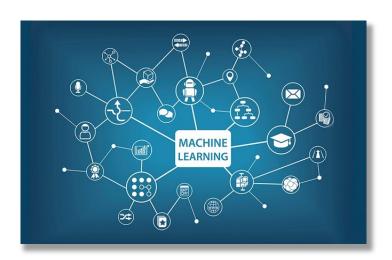


Titan on the Red Project





- ---> Automatically extract metadata from digitized documents
- ---> Perform natural-language search
- ---> Enforce security classification and NTK protocols
- ---> Be expandable to ingest various data stores (Online Vault, PDMLink, shared drives, SharePoint, etc.)
- ---> Utilize commercial, public domain, and LANL ontologies



Core Data Platform:

- ---> Company found through contacts in the US Intelligence Community
- ---> Company worked for several of the IC's three-letter agencies
- ---> Successfully funded (\$800K) and completed a 6-month demo on the unclassified network in FY20 and passed the Red Team evaluation
- ---> Brought capability on-contract for classified network in FY21
- FY21 efforts funded through combination for funds from NA-115 and LANL



LANL Automated Cataloging/Search Project Timeline





2020
Pilot with Titan
Test for functionality
and classified

implementation

2022
Main data
processing
Begin processing main
data sources

↓ 2016Technical evaluation

Evaluation of seven tools

2nd technical evaluation

2018

Targeting AI/ML tools from US intelligence community

Classified environment

2021

Implement on classified systems and infrastructure

Processing and expansion

2023+

Continue to incorporate information from main data sources and evaluate expansion







